



US005335289A

United States Patent [19]

Abdelazim

[11] Patent Number: 5,335,289
[45] Date of Patent: Aug. 2, 1994

[54] RECOGNITION OF CHARACTERS IN CURSIVE SCRIPT

[75] Inventor: Hazem Y. Abdelazim, Cairo, Egypt

[73] Assignee: International Business Machines Corporation, Armonk, N.Y.

[21] Appl. No.: 837,450

[22] Filed: Feb. 14, 1992

[30] Foreign Application Priority Data

Feb. 13, 1991 [EG] Egypt 91301146.6

[51] Int. Cl.⁵ G06K 9/34

[52] U.S. Cl. 382/9; 382/39;
382/14; 382/20

[58] Field of Search 382/9, 13, 36, 39, 40,
382/22, 37, 30, 34, 14, 20

[56] References Cited

U.S. PATENT DOCUMENTS

3,643,215	2/1972	Ingham et al.	382/21
4,288,779	9/1981	Otsu et al.	382/18
5,034,991	7/1991	Hagimae et al.	382/36
5,067,166	11/1991	Ito	382/37
5,075,896	12/1991	Wilcox et al.	382/39

FOREIGN PATENT DOCUMENTS

2516671 6/1974 Fed. Rep. of Germany G06K 9/06

OTHER PUBLICATIONS

H. Y. Abdelazim & M. A. Haskish, "Automatic Reading of Bilingual Typewriter Text", IEEE Confer., Hamburg, GE, vol. 2, pp. 140-144, May 8-12, 1989.

M. Krlmakhem & M. C. Ferri, "Arabic Typewritten Character Recognition Using Dynamic Comparison",

Proc. 1st Kuwait Computer Conf., pp. 455-462, Mar. 1989.

T. Sheikh & Guindi, "Computer Recognition of Arabic Cursive Scripts" Pattern Recognition, vol. 21, No. 4, 1988, pp. 293-302.

R. Casey et al., "Recursive Segmentation & Classification of Composite Character Patterns", IEEE Computer Society Pres., pp. 1023-1026, Oct. 19-22, 1982.

T. Shimizu, "Pattern Recognition Device", Patent Abstracts of Japan, vol. 12, No. 468, p. 797, Dec. 8, 1988.

Bozinovic et al. "Off-Line Cursive Script Word Recognition" IEEE Transactions on Pattern Analysis and Machine Intelligence, VII #1 Jan. 1989, pp. 68-83.

Primary Examiner—Leo H. Boudreau

Assistant Examiner—Larry J. Prikockis

Attorney, Agent, or Firm—Lauren C. Bruzzzone; Joseph C. Redmond, Jr.

[57] ABSTRACT

A system and method for recognizing characters in cursive script is provided in which the script is scanned to detect word boundaries and words are then segmented into characters. This is accomplished by segmenting the script to form an initial portion, the segmentation being performed with reference to its position relative to a word boundary. This initial portion is then compared with a set of reference portions. Subsequent portions of the script are taken in sequence and compared with reference portions until a character is identified with an uncertainty less than a predetermined threshold value. A new initial portion is then segmented, with the new initial portion chosen on the basis of the average width of the character identified and the comparison process repeated to identify the next character.

11 Claims, 3 Drawing Sheets

